

Scientific Data Management at European XFEL.



Nasser Alqudami¹, Steve Aplin¹, Djelloul Boukhelef¹, Fabio Dall'Antonia¹, Illia Derevianko¹, Uwe Ensslin², Martin Gasthuber², Juergen Hannappel², Mwai Karimi², Thomas Kluyver¹, Luis Maia¹, Janusz Malka¹, Maurizio Manetti¹, Tigran Mkrtchyan², Kars Ohrenberg², Carsten Patzke², Gianpiero Previtali¹, Philipp Schmidt¹, Kilian Schwarz², Egor Sobolev¹, Janusz Szuba¹, Bharathi Vanganuru¹, Christian Voss², Krzysztof Wrona¹, Christopher Youngman¹

¹European XFEL GmbH, ²Deutsches Elektronen-Synchrotron DESY

Abstract / Introduction

Data Management is essential to make research data easily accessible and usable. Important ingredients of data management include data policies and data workflows. The data workflows are based on the policies which are implemented by defining a set of parameters stored in the metadata catalogue. The role of the metadata catalogue in relation to the data management services and underlying hardware solutions for the data storage systems will be presented. The architecture of the storage system consists of four layers, each addressing a different set of challenges. The first – online - is designed as a fast cache for the data generated directly at the scientific instruments during experiments. The second layer – offline - provides the performance for data processing during and after the beamtimes. The third layer - dCache disk pool - delivers the capacity to the system for long-term storage and the last one - tape archive - provides data safety and long-term archive. The storage system is able to accept 2PB/day of raw data, demonstrating the real capabilities with all sub-services being involved in this process. The storage system is connected to the high-performance computing cluster supporting remote data analysis and alternatively allows external users to export data outside of the European XFEL facility.

Essential ingredients of Scientific Data Management

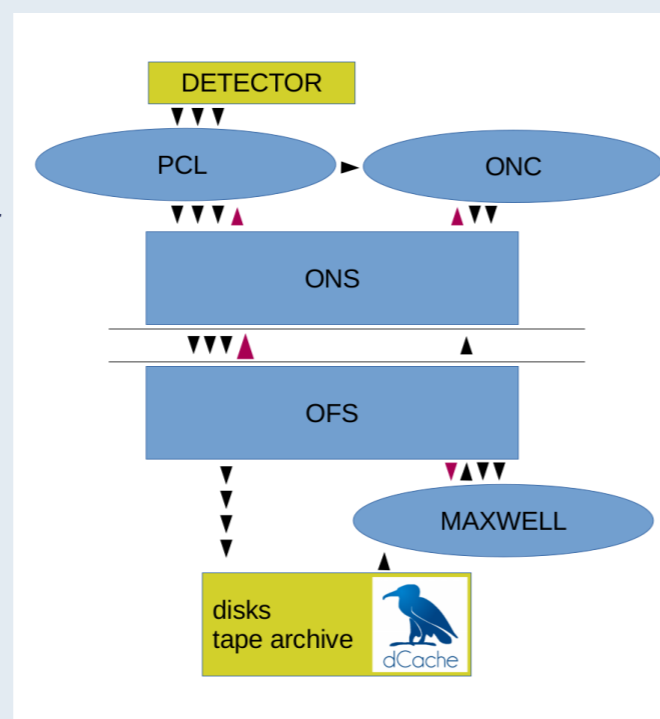
- Scientific Data Policy:
 - Defines the rules for scientific data usage
 - Based on PaN-data Europe Strategic Working Group established in 2011
 - Each user accepts it upon registration in the European XFEL user portal (UPEX)
 - Scientific Data Policy has been efficiently used as a base for implementation of data management services
 - Planned update of the Scientific Data Policy according to the FAIR principles



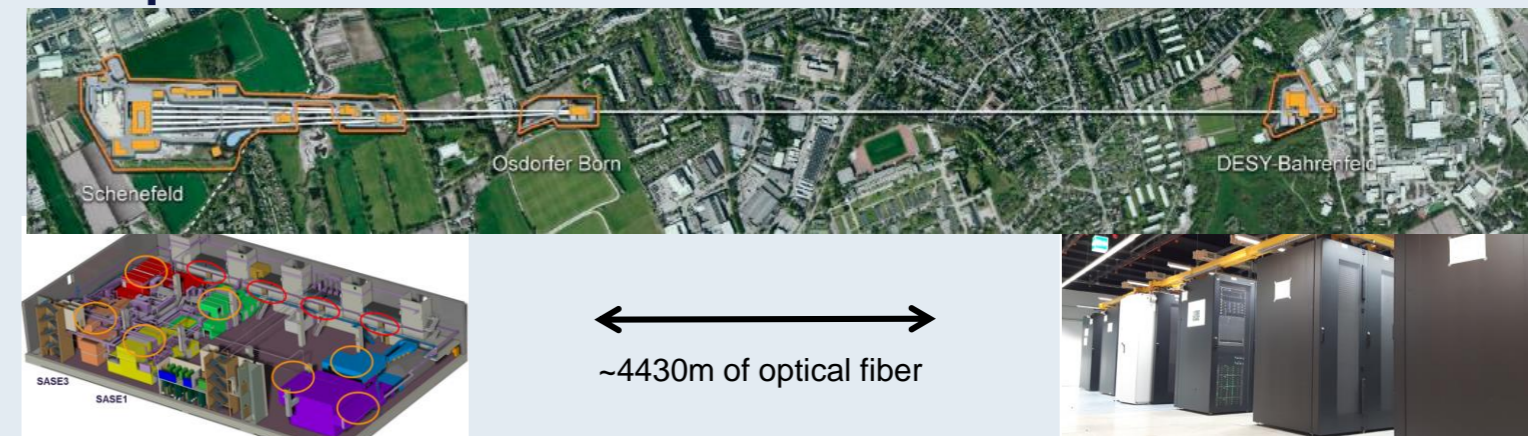
- Data management services implementing defined policies:
 - User portal (UPEX) - proposals handling, samples, scheduling and organization of beamtimes
 - Metadata Catalogue and Data Management Portal (myMDC) - data acquisition, access control, data migration and export, control of data calibration and processing
 - Underlying Data Infrastructure – Computing, Storage, Services

High performance Scientific Data Infrastructure

- PC-Layer (DAQ) infrastructure
 - Dedicated resources per SASE tunnel
 - Data aggregation from distributed data sources (detectors, digitizers, sensors, ...)
 - Data formatting and HDF5 file creation
 - Streaming data to on-line computing cluster
- ONC – online computing cluster
 - Dedicated resources per SASE tunnel
 - Captures raw data stream from PCL-Layer
 - Initial on-line calibration of raw data
 - On-line data preview and analysis
- Online Storage (ONS) - Cache
 - High performance
 - Optimized for concurrency
 - High redundancy
 - Dedicated storage per SASE tunnel
 - Capacity for a few days
- Offline Storage (OFS)
 - High performance storage
 - Shared data storage for all SASEs
 - High redundancy
 - Data source for the off-line data analysis
- Offline computing cluster (MAXWELL)
 - Shared among facilities
 - Generation of calibration constants
 - Calibrating and processing of Raw data
 - End User data analysis
- dCache
 - High capacity
 - Primary raw data source after data retention
 - Long term storage for the raw data
 - Interface to the tape archive (second copy)



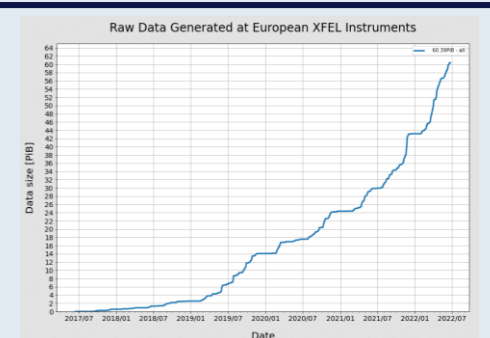
European XFEL - overview



Three computer (balcony, red) rooms in the experiment hall EuXFEL hardware in DESY Computer Center

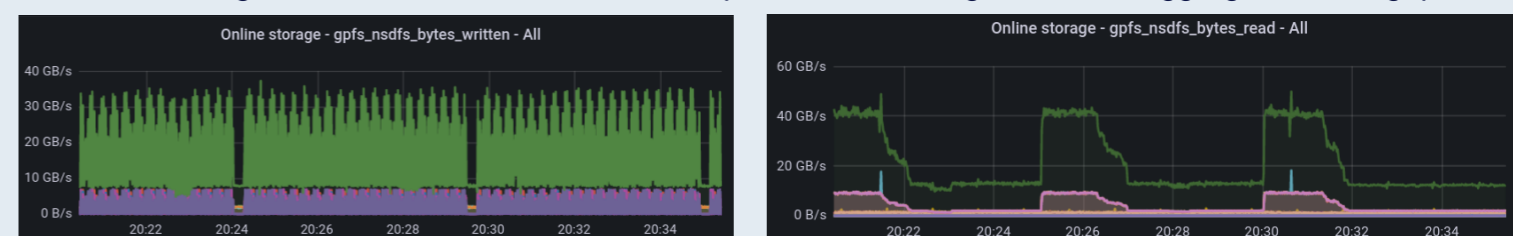
Increased data taking efficiency

- The average effective time of data taking grows
- Size of raw data collected in 2021 RAW: 19PiB
- Size of raw data collected in first half of 2022: 18PiB
- Increased size of corrected/calibrated data generated by the facility provided calibration pipeline
- Corrected and RAW size might reach 70PiB in 2022
- New techniques of data handling are required to keep up with the increased data rates
- Data selection, shorter retention periods, data processing reproducibility

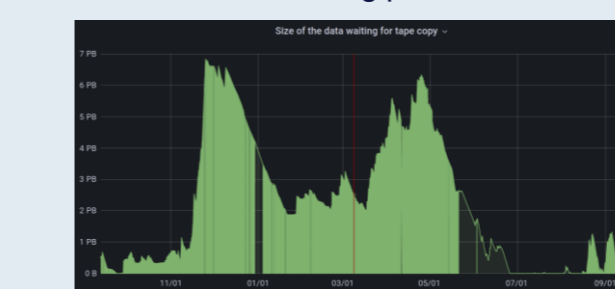


Performance

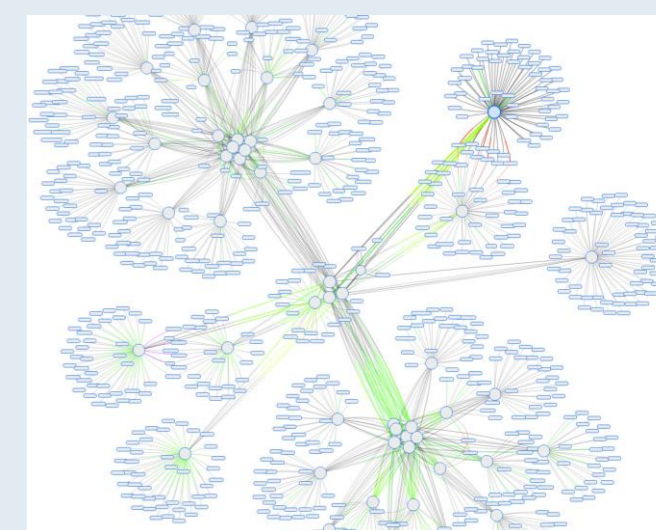
- Online storage
 - Infrastructure based on IBM Elastic Storage Server (3PB + 2x1.5PB)
 - Up to 60GB/s write performance
 - Accepts 15GB/s data ingest from 2D detectors (AGIPD, LPD, DSSC)
 - Data drain between SSD and HDD layer (up to 15GB/s)
 - Data migration to offline cluster of 30GB/s per SASE, scaling to 90GB/s aggregated throughput



- Offline storage
 - Connected to online systems with 1Tb/s link provided by HDR IB switches
 - 40PB of disk storage
 - Theoretical maximum performance of 300GB/s
 - Observed reading performance of about 180GB/s



- dCache
 - 120PB of disk storage
 - 2PB/week copy to tape performance



Summary / Outlook

- European XFEL supports the full scientific data life cycle from the experiment proposal submission through the data generation during the beamtime and further remote analysis up to the long term data archive.
- Implemented data management system has demonstrated the capability to cope with the extremely high data throughput and data volume requirements
- Update of Scientific Data Policy is required to meet the FAIR data principles
- Update of data retention policies and efficient data reduction techniques are indispensable to sustain the observed exponential grow of data storage requirements

Acknowledgements

We wish to acknowledge the help provided by the instrument scientists and data experts of European XFEL GmbH and support of our DESY-IT colleagues not mentioned in the author lists. We would also like to show our deep appreciation to our business partners: IBM and Nvidia which are helping us to provide an excellent data service for users of our facility.